Observatoire Economique et Statistique d'Afrique Subsaharienne

Atelier régional des Nations unies sur le traitement des données du recensement : les technologies contemporaines pour la saisie et correction des données du recensement, Bamako, Mali
3-7 novembre 2008

Activités d'AFRISTAT dans le cadre des recensements de la population et de l'habitat

&

Introduction au choix des méthodes de saisie et traitement des données

Poirel Guillaume
Expert Démographe – AFRISTAT

SOMMAIRE

I. HISTORIQUE ET AXES DE TRAVAIL D'AFRISTAT 1

II. LES STATISTIQUES DEMOGRAPHIQUES ET SOCIALES DANS LE PSTA 2

II.1 Les activités d'AFRISTAT en matière de recensement *2*

II.1.1 Réflexion sur l'optimisation du RGPH 3
II.1.2 Réflexion sur l'optimisation de la cartographie censitaire 4
II.1.3 Réflexion sur la mise en œuvre d'un recensement de la population et de l'habitat simultané et harmonisé sur les pays de l'UEMOA 4

II.2 Les activités d'AFRISTAT en matière d'état civil *4*

III. POURSUITE DES ACTIVITES D'AFRISTAT 5

III.1 Un calendrier à prendre en considération *5*

III.2 Un renforcement des partenariats *5*

III.3 La poursuite des reflexions et l'organisation d'un séminiare international *5*

III.4 Répondre aux demandes d'appuis des pays *6*

INTRODUCTION AUX METHODES DE TRAITEMENTS DES RECENSEMENTS **7**

I. METHODE DE TRAITEMENT OU DE SAISIE 7

I.1 Avantages et inconvénients de chaque méthode *7*

I.2 Recommandations liées aux méthodes de saisie *11*

II. METHODE DE CODAGE 11

II.1 II.1 Avantages et inconvénients de chaque méthode *12*

II.2 II.2 Recommandations liées aux méthodes de codage *12*

III. METHODE D'APUREMENT DES DONNEES 12

III.1 Avantages et inconvénients de chaque méthode *14*

III.2 Recommandations liées aux méthodes de correction *14*

IV. AUTRES ASPECTS DU TRAITEMENT DES DONNEES 15

IV.1 Stockage et archivage *15*

IV.2 Recommandations liées aux méthodes de correction *15*

CONCLUSION **15**

PRESENTATION D'AFRISTAT ET DE SES ACTIONS

I. HISTORIQUE ET AXES DE TRAVAIL D'AFRISTAT

L'Observatoire Economique et Statistique d'Afrique subsaharienne (AFRISTAT) est un outil de renforcement des capacités statistiques et d'harmonisation des concepts et méthodes nécessaire à l'exécution et au suivi de l'intégration économique régionale. Crée en 1993 par 14 pays membres de la Zone Franc pour répondre efficacement aux nouveaux défis statistiques, il regroupe aujourd'hui dix huit (18) pays[1]. Son objectif général est de « contribuer à l'élaboration, à la mise en œuvre et au suivi et évaluation des politiques de développement de la statistique des Etats d'Afrique subsaharienne ». AFRISTAT intervient sur le renforcement des capacités des systèmes statistiques nationaux (SSN) et les travaux d'harmonisation des données statistiques nécessaires aux politiques d'intégration sous-régionale en participant à :

- La construction des outils de base (socle minimal) pour le développement de la statistique ;
- au développement des systèmes statistiques cohérents et efficaces ;
- la formulation de réponses aux demandes d'information pour la conception, la mise en œuvre et le suivi et évaluation des politiques de développement ;
- la formation des cadres nationaux en matière de traitement et de diffusion de l'information statistique.

Dans les premières années de son existence et pour atteindre ses objectifs, AFRISTAT s'est principalement consacré à l'harmonisation des méthodologies de calcul des indices de prix, des nomenclatures statistiques et des méthodes d'élaboration des comptes nationaux. Ensuite, AFRISTAT s'est doté en septembre 2001 et pour la période 2001-2005 d'un ***Programme Statistique Minimum Commun*** (PROSMIC) qui a mis l'accent sur la réalisation d'actions de renforcement de capacités des Etats membres dans les domaines de la coordination statistique, la comptabilité nationale, la conjoncture économique et sociale, la diffusion, les conditions de vie des ménages et les statistiques agricoles.

Depuis décembre 2005, le ***Programme Stratégique de travail d'AFRISTAT*** (PSTA) pour la période 2006-2010 prend en compte les besoins formulés par les pays dans le domaine des statistiques sociales et de la démographie. Ainsi le PSTA comporte six volets :

i) appui au renforcement des capacités institutionnelles des Etats membres en matière d'organisation des systèmes statistiques nationaux ;

ii) appui au développement des systèmes d'information pour la conception, la mise en œuvre et le suivi et l'évaluation des stratégies de réduction de la pauvreté ainsi que pour la mise en œuvre des OMD ;

[1] Bénin, Burkina-Faso, Burundi, Cameroun, Centrafrique, Comores, Congo, Côte d'Ivoire, Gabon, Guinée, Guinée Bissau, Guinée Equatoriale, Mali, Mauritanie, Niger, Sénégal, Tchad, Togo,

iii) appui au renforcement des capacités de collecte, de traitement et d'analyse des données statistiques économiques et d'élaboration de synthèses statistiques et économiques ;

iv) assistance technique aux SSN dans le développement de bases de données et de diffusion statistique ainsi que la mise à leur disposition de méthodes et outils statistiques ;

v) recherche appliquée en vue du développement des méthodologies adaptées aux capacités des systèmes statistiques nationaux ;

vi) renforcement des moyens de travail d'AFRISTAT.

Inscrivant ses interventions dans ces différents volets, AFRISTAT exécute également ***d'autres programmes et projets*** à la demande des partenaires techniques et financiers parmi lesquels :

- des travaux d'harmonisation des comptes nationaux en vue d'améliorer la comparabilité des agrégats servant à déterminer les critères de convergence. Les travaux d'harmonisation de comptes nationaux entrepris pour l'UEMOA ont été étendus à la CEDEAO avec l'appui d'AFRISTAT ;
- dans le même esprit, des projets statistiques ont été (et sont) menés dans le cadre de l'harmonisation des indices des prix à la consommation dont la mise en œuvre a été confiée à AFRISTAT (1996, 2007) ;
- les activités réalisées dans le cadre du Programme d'appui régional à la statistique (PARSTAT) (production et harmonisation des informations en matière d'Indice des Prix à la Consommation, de Comptes Nationaux, d'études du Secteur Informel) ;
- le projet d'appui au suivi et évaluation des DSRP et des OMD dans une vingtaine de pays africain avec le PNUD ;
- Le projet IDF de la Banque Mondiale sur l'archivage des données ;
- le projet pilote « d'amélioration de la qualité des statistiques du marché du travail et le renforcement de la gestion de l'information du marché du travail et des systèmes de suivi de la pauvreté en AFRIQUE » avec ACBF et le BIT ;
- le suivi de la mise en œuvre du Programme de comparaison international (PCI) avec la BAD.

II. Les statistiques démographiques et sociales dans le PSTA

II.1 LES ACTIVITES D'AFRISTAT EN MATIERE DE RECENSEMENT

Pour tenir compte de ces sollicitations dans le domaine sociodémographique, l'objectif 5 du volet 2 du PSTA prévoit de « participer à l'élaboration des méthodologies peu coûteuses de collecte et de traitement de données démographiques par des enquêtes et recensements, y compris l'état civil, et aider les Etats membres à les mettre en œuvre ».

Les activités liées aux appuis apportés aux Etats dans la préparation des recensements sont les suivantes :

i) participation aux réunions et forum nationaux et internationaux sur la préparation de la série des recensements 2010 ;

ii) vulgarisation des manuels et des guides harmonisés sur les recensements de la population ;
iii) organisation des séminaires et ateliers sur la pratique des recensements de la population ;
iv) appui technique à apporter aux Etats pour l'organisation des recensements de la population.

Bien au-delà de ces activités, plusieurs réflexions en matière de recensement ont été initiées par AFRISTAT. La spécificité du recensement général de population et de l'habitat réside dans le caractère exhaustif de ses opérations d'observation de la population. Elle conforte les trois objectifs qui lui sont assignés à savoir : i) le dénombrement de la population au niveau géographique le plus fin ; ii) le recueil des informations sociodémographiques ; iii) et la constitution d'une base de sondage. AFRISTAT est particulièrement sensible aux deux derniers objectifs qui établissent le lien entre le recensement de la population et le système d'enquêtes auprès des ménages. AFRISTAT travaillant sur les enquêtes ménages des pays membres sait exactement quels sont les besoins en matière de bases stratifiées dont il aura bientôt besoins pour lancer les enquêtes 1-2-3 national avec un panel emploi par exemple.

II.1.1 REFLEXION SUR L'OPTIMISATION DU RGPH

- Publication d'un article sur « l'optimisation des recensements en Afrique subsaharienne », La Lettre d'AFRISTAT numéro 40, Janvier 2007 ;
- AFRISTAT, « National Strategy of development of statistic : lever for the restoration of the censuses in sub-Saharan Africa », Kigali, Rwanda – January 15 -19 2007;
- AFRISTAT, « Optimisation du RGPH et de ses prolongements vers les secteurs sociaux et le suivi-évaluation des SRP », 12ème session du Conseil Scientifique d'AFRISTAT, Bamako du 24-26 septembre 2007.

Cette réflexion a permis de mettre en avant les réelles menaces sur les recensements en Afrique subsaharienne et la nécessité d'une approche alternative adaptée, reposant sur une adéquation optimale des moyens aux fins. Cette réflexion apparaît comme un dispositif de sensibilisation pour le financement de l'opération. Plus explicitement, il s'agit d'une démarche d'optimisation des coûts, d'optimisation de l'information, faisant appel à tous les acteurs, en tirant partie de toutes les synergies possibles et se référant à une cohérence d'ensemble. La stratégie d'optimisation du recensement devrait porter, de manière complémentaire, sur les éléments moteurs suivants :

- la simplification des questionnaires et l'optimisation de l'information ;
- l'utilisation des nouveaux outils technologiques (SIG, traitement, diffusion) ;
- le développement des sources administratives ;
- la mise en place d'une stratégie de développement statistique.

Etant donné la complexité des éléments techniques, administratifs et environnementaux en jeu, il ne peut exister une proposition unique pour optimiser les recensements. Il s'agit donc en fonction de la situation de chaque pays, de rechercher l'option qui allie au mieux les objectifs, les outils méthodologiques, les ressources et les produits.

II.1.2 REFLEXION SUR L'OPTIMISATION DE LA CARTOGRAPHIE CENSITAIRE

- AFRISTAT, « Cadre d'optimisation de la phase de cartographie censitaire », Atelier régional des Nations Unies sur l'organisation et la cartographie des recensements, Rabat, Maroc, 12-16 novembre 2007.

En effectuant l'état des lieux du secteur de la cartographie national au Mali (producteurs, utilisateurs), l'étude montre qu'il existe de plus en plus d'utilisateurs de SIG et de producteurs d'informations cartographiques. L'optimisation passe par l'utilisation partagée des produits issues de l'ensemble des acteurs de la cartographie afin d'alléger l'opération de cartographie censitaire (coûts humains, logistiques) et d'améliorer les résultats en continue (plus de diffusion). Dans le cas étudié, l'optimisation se réalise à travers la mise en commun des informations cartographiques par : i) l'amélioration des géoréférencement et du bornage ; ii) le géocodage unique ; iii) l'amélioration des métadonnées et la coordination logistique ; iv) la mise en œuvre d'un partenariat des moyens.

L'étude explique que cette démarche ne peut se réaliser sans un renforcement de la coordination et la mise en place de partenariats par la création ou dynamisation d'un Comité National de l'Information géographique (réseau d'utilisateur), la création d'une structure de centralisation des données cartographiques des partenaires (producteurs/utilisateurs), la diffusion des activités cartographique de l'institut de statistique durant la période intercensitaire ou encore la mise en commun sur un serveur.

II.1.3 REFLEXION SUR LA MISE EN ŒUVRE D'UN RECENSEMENT DE LA POPULATION ET DE L'HABITAT SIMULTANE ET HARMONISE SUR LES PAYS DE L'UEMOA

AFRISTAT a entamé un réflexion innovatrice en donnant un aperçu des actions possibles au niveau sous régionale pour préparer un recensement de la population et de l'habitat simultané et harmonisé dans les pays de l'Union Economique et Monétaire Ouest Africaine. L'idée serait d'amener les pays de l'UEMOA à effectuer leur RGPH à la même période selon la même méthodologie et des supports de collecte identiques afin de faciliter les comparaisons dans la sous région et en répondant au contexte particulier de l'intégration régionale. L'analyse met l'accent sur certains aspects de ce que pourrait être le RGPH simultané et harmonisé en mettant l'accent sur les justifications de l'approche, son intérêt et les gains potentiels, et enfin quelques aspects méthodologiques (Conseil scientifique d'AFRISTAT).

II.2 LES ACTIVITES D'AFRISTAT EN MATIERE D'ETAT CIVIL

AFRISTAT a également effectué des appuis aux Etats membres (Sénégal, Gabon, Niger) pour la relance de la collecte et de l'exploitation des données de l'état civil. Si cet aspect est abordé ici, c'est que le système d'information démographique dont le pilier majeur reste le recensement, comprend d'autres sources telles que les enquêtes démographiques ou l'état civil. L'amélioration des sources administratives est l'une des opportunités pour permettre l'amélioration du système d'information démographique et l'appropriation de nouvelles méthodes de dénombrement. Tout comme les statistiques de l'éducation ou de la santé, le lancement de l'exploitation des statistiques de l'état civil connaît des débuts difficiles, mais

pourrait amplement participer à l'amélioration de l'enregistrement à travers la diffusion des résultats obtenus (sensibilisation), la comparaison avec d'autres sources. D'ailleurs le taux d'enregistrement des naissances fait parti des indicateurs de bonne gouvernance du DSRP de la Guinée. En travaillant sur l'état civil, AFRISTAT participe également au renforcement du système d'information démographique sur le long terme.

III. Poursuite des activités d'AFRISTAT

III.1 UN CALENDRIER A PRENDRE EN CONSIDERATION

AFRISTAT fonctionne en répondant aux demandes des pays membres et jusqu'à ce jour ces dernières n'ont pas été nombreuses. Mais le calendrier des recensement des pays membres d'AFRISTAT (Annexe1) indique clairement les demandes risquent d'être nombreuses dans les deux années à venir. Certains pays ont effectué l'opération il y a peu de temps (Burkina Faso, Congo) mais la majorité sont en cours de préparation (Mali, Guinée Bissau, Côte d'Ivoire, Guinée équatoriale, Guinée Conakry, Tchad, Togo) ou en plein dans l'opération (Burundi) voir en phase post dénombrement.

III.2 UN RENFORCEMENT DES PARTENARIATS

L'introduction récente des questions de population dans le programme de travail d'AFRISTAT exige un renforcement des partenariats avec les partenaires technique et financier (UNFPA, DSNU, IFORD, INSEE, EUROSTAT, CEPED…) mais également avec les pays. La participation d'AFRISTAT au Programme mondial sur les recensements de population et de l'habitat découle de ses missions et procède de la coopération technique en matière de statistique. Le Programme mondial sur les recensements de population en formulant les recommandations des recensements de population participe à l'harmonisation des données, des définitions et de la classification des caractéristiques. Il améliore ainsi la comparabilité internationale des informations censitaires. AFRISTAT accompagnera les pays à la mise en œuvre de ces principes et recommandations. Le processus d'harmonisation des concepts et définitions dans la sous-région est un processus à long terme qui nécessite de prendre des précautions : l'exemple du concept de résidence en est un exemple (Cf. annexe 2).

III.3 LA POURSUITE DES REFLEXIONS ET L'ORGANISATION D'UN SEMINIARE INTERNATIONAL

De part sa connaissance approfondie des systèmes statistiques nationaux (SSN) et des INS des pays membres, AFRISTAT a une vision claire et suivie des réalités et difficultés que rencontrent les pays. Il est évident que l'application des principes et recommandations doit s'effectuer en prenant en compte les spécificités et contextes de la sous région. L'ensemble des réflexions seront poursuivies et proposées afin de sensibiliser et d'accompagner les pays dans la mise en œuvre de leur recensement.

Enfin, selon le relevé de conclusion de la 18^{ième} réunion du Comité de Direction d'AFRISTAT qui a eu lieu les 20 et 21 octobre 2008 à Bamako, AFRISTAT prévoit d'organiser en marge

de la 19ième réunion du Comité d'AFRISTAT avec de nombreux partenaires un séminaire international sur les recensement et la valorisation des données démographiques.

« 12. Le Comité de direction a retenu la proposition d'organiser en octobre 2009 un séminaire international sur les recensements de la population et de l'habitat et la valorisation des données démographiques en collaboration avec l'IFORD et le CEPED. Il invite la Direction générale à rechercher les financements avec l'appui des partenaires techniques et financiers. ». (Extrait du relevé de conclusions de la 18ème réunion du Comité de direction d'AFRISTAT).

III.4 REPONDRE AUX DEMANDES D'APPUIS DES PAYS

AFRISTAT continuera à répondre aux demandes d'appuis techniques et demandes d'accompagnement des pays dans le cadre de l'ensemble du processus de préparation, de mise en œuvre et de valorisation de l'opération. Pour cela, les pays sont sollicités à établir des demandes clairement explicitées pour qu'AFRISTAT puisse y réponde le plus rapidement possible. AFRISTAT dispose de certains atouts :

- connaissance de l'organisation statistique dans les Etats membres ;
- contribution au renforcement des capacités statistiques des systèmes statistiques nationaux ;
- est connu et reconnu par la plupart des parties prenantes ;

Dans le cadre des recensements africains AFRISTAT peut, notamment :

- apporter une expertise technique aux pays selon les domaines d'expertise d'AFRISTAT ;
- apporter son appui technique aux Etats notamment dans l'amélioration des outils de collecte, de l'utilisation des données de la cartographie, de l'archivage, de la diffusion, etc. ;
- faciliter la vulgarisation des manuels et des guides harmonisés sur les recensements ;
- contribuer à rendre plus probable la mise en œuvre des recommandations par l'appui aux pays désirant qu'AFRISTAT participe à l'élaboration des documents, à la préparation des activités du RGPH ;
- Accompagner les pays dans la préparation de leur RGPH sur les différentes phases de l'opération ;
- contribuer à l'amélioration des statistiques administratives (santé, éducation, état civil) notamment l'amélioration de leur couverture

Pour cela AFRISTAT compte proposer un accompagnement à la DNSI du Mali afin de valider ces réflexions et être à même d'avoir une vision encore plus exacte des difficultés rencontrées.

INTRODUCTION AUX METHODES DE TRAITEMENTS DES RECENSEMENTS

I. Méthode de traitement ou de saisie

Le choix de la méthode de traitement (saisie des données) a un impact direct sur le succès global du recensement. Les progrès techniques réalisés permettent aujourd'hui aux instituts en charge du RGPH de traiter et de saisir les données de façon plus précise. D'ailleurs les principes et recommandations des Nations Unies sensibilisent les pays à utiliser ces nouvelles technologies qui possèdent certains avantages tels que le gain en temps ou encore une diffusion plus rapide des résultats et à des échelles plus fines. Ainsi, le travail d'exploitation est devenu de plus en plus efficace et économique. Les nouvelles technologies ont permis une amélioration des produits diffusés et des résultats offerts aux utilisateurs.

AFRISTAT pourra accompagner les pays désirant utiliser ces techniques de plus en plus répandues (imagerie, reconnaissance des caractères, lecture optique) en analysant leurs caractéristiques, les gains, les obstacles et les risques. En effet, le choix de la méthode de traitement des données doit prendre en compte plusieurs aspects :

1. Pérennité du système (pour combien de temps)
2. Délais de démarrage (s'y prendre en avance)
3. Infrastructures technologiques nécessaires
4. Moyens humains et qualification
5. Coûts financiers

Ces aspects sont évoqués à travers l'approche optimisée du recensement qu'AFRISTAT a entamé depuis plus d'une année. Le choix entre les méthodologies dites manuelles et automatiques devrait résulter d'une analyse du coût du personnel et du coût du matériel et d'une analyse des délais à respecter et des résultats à diffuser. D'autre part, l'atelier permettra également d'aborder la question de la sous-traitance de certaines opérations.

I.1 AVANTAGES ET INCONVENIENTS DE CHAQUE METHODE

Le tableau 1 présente de façon générale les méthodes de traitement et de saisie, leurs avantages et inconvénients. Les expériences pays permettront de rentrer dans les détails et de voir dans quel contexte chaque méthode est viable ou possible à mettre en œuvre (faisabilité). Mais il faut garder tout de même à l'esprit que la simulation du processus de fonctionnement du cerveau humain que l'on rencontre dans les technologies avancée est complexe. L'un des objectifs pour chaque responsable des opérations de recensement reste toujours de faciliter chaque opération (saisie par exemple) en fonction du contexte et en évitant d'introduire de nouvelles difficultés qui risqueraient de retarder l'ensemble de l'opération et la publication des résultats.

TABLEAU 1. METHODES DE SAISIE ET TRAITEMENT

METHODES	AVANTAGES	INCONVENIENTS	OBSERVATIONS/REMARQUES
Saisie au Clavier	Logiciel Simple	Personnel nombreux	Cette méthode de traitement est préconisée lorsque les infrastructures informatiques sont peu importantes et que le coût du personnel est faible.
		Temps du traitement important (délais long)	
		Nécessite de regrouper les variables descriptives par catégorie et procéder à leur codification et classification	
		Matériel de calcul de faible capacité	
OMR - Lecture Optique de marques	Plus rapide que la saisie manuelle	Nécessite que certaines réponses soient saisies manuellement ou codées	Le coût de cette méthode est plus important à cause de la précision de l'impression, de l'utilisation exclusive d'un type de papier, encre format. Cette méthode est préconisée lorsqu'il existe une rigueur des agents recenseur.
	Coût raisonnable	Nécessite une précision de remplissage des questionnaires (stylo ou crayon adéquate) et d'impression standard	
	Simplicité de sa mise en place	Nécessite un type de papier et d'encre adéquate	
	Utilisation peu difficile	Restriction sur le format du questionnaire	
	De nombreuses expériences ont approuvé la méthode		
Imagerie numérique (scanner, caméra numérique)	Permet une sauvegarde de chaque questionnaire	S'associe à la lecture optique et l'exploitation des questionnaires et se situe juste à la marge de la reconnaissance intelligente des caractères	Le résultat est la production d'une image et donc d'un fichier de données. La saisie (dépouillement) est effectuée.
	Logiciel de reconnaissance des cases cochées (lecture optique de marque)		
	Permet les économies de papier (les questionnaires sont remplacés par des images)		
	Permet une grande productivité (car plus efficace)		

METHODES	AVANTAGES	INCONVENIENTS	OBSERVATIONS/REMARQUES
ICR - Reconnaissance intelligente de caractères (scanner)	Permet une sauvegarde de chaque questionnaire	Nécessite la création de lexiques de variables pour les corrections potentielles	Logiciel de reconnaissance des cases cochées (lecture optique de marque) et des notations écrites dans un espace limité. Cette méthodologie permet la transcription des notations écrites en caractères, une correction plus ou moins automatique et un codage automatique. Il s'agit d'une nouvelle technologie qui comporte des risques, raison pour laquelle il est nécessaire d'effectuer des essais rigoureux.
	Permet d'effectuer des économies sur :	Nécessite quand même des opérateurs pour les corrections	
	1. les salaires (peu de personnel pour le codage)	Coût élevé des équipements dû à la complexité du matériel et du logiciel	
	2. la manutention matérielle et loyers	Volume de données important (coût matériel) et capacité de stockage important	
	3. les dépenses de personnel (procédures de recrutement, soutien administratif)	Substitution possible de caractère pouvant nuire à la qualité des données (sur des variables importantes) selon la qualité du logiciel et du dispositif de reconnaissance	
	Permet une augmentation les rendements	Nécessite d'effectuer des réglages particuliers et délicats du dispositif de reconnaissance pour les seuils de confiance. Ceci a un impact sur la qualité et le coût de l'activité	
	Permet une amélioration de la qualité des données (codage automatique) et meilleure cohérence du traitement des réponses	Obligation de respecter des règles d'écritures (espace alloué à chaque lettre par exemple)	
	Permet de réduire les délais de saisie et d'apurement et donc les délais de publication	Le codage automatique peut nuire à la qualité des données (dans le cas d'erreurs systématiques)	
	Ne nécessite pas un format type d'impression comme la lecture optique	Les économies peuvent ne pas être effectuées dans le cas où le codage est plus lent que prévu	
	Permet une grande précision des corrections malgré l'état du questionnaire	Peu d'expérience et donc risques importants	

METHODES	AVANTAGES	INCONVENIENTS	OBSERVATIONS/REMARQUES
Questionnaires électroniques	Réduction des coûts humains liés à la saisie	L'ensemble des ménages doit disposer d'un ordinateur et d'un accès à Internet (taux d'accès au Mali)	Il s'agit de recourir à un Internet avec des systèmes de saisie automatique des informations. Cette méthode permet de réduire les coûts humains nécessaire à la saisie au traitement, au codage des questionnaires.
	Réduction des coûts humains liés au traitement	Difficulté de déterminer le double compte et l'exhaustivité	
	Réduction des coûts humains liés au codage	Nécessité une sécurisation pour la transmission des informations	
		Nécessité des systèmes de collecte parallèles et engendre des difficultés de fusion des données de différentes origine	
Interview avec ordinateur portable		Coût très élevé	AFRISTAT ne recommande pas l'utilisation de ces technologies dans la plupart des pays d'Afrique subsaharienne. En effet, ces technologies nécessitent des pré requis importants (environnement, poussière etc.). Cette méthodologie est actuellement peu viable.
		Nécessite de disposer de ressources humaines qualifiées nombreuses et rigoureuses	

Bien que les nouvelles technologies de saisie offre des possibilités de gains en temps, en coûts ou en qualité, les pays de la région rencontrent des spécificités et un contexte qui nécessitent de nombreuses précautions. La méthode de reconnaissance intelligente des caractères (ICR) nécessite une gestion accrue des risques en :

1. réduisant au maximum les espaces réservés pour la reconnaissance des caractères ;
2. évitant la substitution pour des variables importantes (également pour minimiser les vérifications par des opérateurs) en maximisant le nombre de variables à cocher ;
3. utilisant au maximum la lecture optique de marques et utiliser uniquement la reconnaissance des caractères pour des modalités « autres » des variables ;
4. utilisant dans la mesure du possible un dispositif double de reconnaissance pour réduire la substitution (exemple variable âge et variable date de naissance) ;
5. utilisant des seuils de confiance élevés pour les variables importantes ou faire des vérifications images par des opérateurs ;
6. assurant une bonne qualité des images ;
7. effectuant un contrôle de qualité par sondage manuel des résultats de chaque opération (mesure de la proportion d'erreur).

I.2 RECOMMANDATIONS LIEES AUX METHODES DE SAISIE

- analyser les inconvénients et les avantages de chaque méthode de saisie, les risques et les gains en fonction du contexte national pour déterminer le choix de la méthode la plus adéquate ;
- effectuer une analyse de faisabilité pour les méthodes de saisie disponible en préconisant une méthode participant au renforcement des capacités et à l'amélioration des compétences ;
- établir un compromis entre la qualité et le coût à travers l'évaluation des seuils de confiance et de nombreux essais lors de l'utilisation de la méthode ICR. Ceci permet d'ailleurs de déterminer un rapport entre la qualité (mesure de l'exactitude) et le coût (correction) ;
- assurer un contrôle de qualité continu pour ajuster sans cesse les seuils de confiance afin d'équilibrer le travail de correction (choix entre codage automatique et manuel) lors de l'utilisation de la méthode ICR ;
- prévoir une capacité de stockage importante et de sauvegarde (fichiers images) pour les méthodes ICR et/ou OMR.

II. Méthode de Codage

Il existe différentes méthodologies de codage (manuel, automatique). Ces dernières évoluent rapidement. Quelque soit la méthode, les activités préparatoires du recensement comprennent l'élaboration des listes ou index des réponses susceptibles d'être formulées par les ménages ainsi que l'élaboration des dictionnaires de codification. Cet outil reste commun à l'ensemble des méthodologies de codage et la mise en œuvre de cette activité à

une conséquence directe sur la qualité des données[2]. Le tableau 2 ci-dessous rappelle les différents types de codage.

TABLEAU 2 : TYPE DE CODAGE	
Codage simple	Il s'agit du codage des variables limitées à une question (lieu de naissance) ou à un nombre de mot limité. Le codage est donc simple car les modalités des variables sont codées facilement
Codage structuré	On appelle codage structuré le codage des variables complexes (activités professionnelles). Il est nécessaire d'utiliser parfois d'autres modalités de variables pour pouvoir effectuer le codage. Ainsi, les procédures et instructions de codage devront prévoir des règles de codage
Codage hiérarchique	Le codage hiérarchie ou successif correspond au codage d'informations nécessitant plusieurs variables (codage d'adresse par exemple)

II.1 II.1 AVANTAGES ET INCONVENIENTS DE CHAQUE METHODE

Quelque soit le type de codage, il existe plusieurs méthodes de codage (tableau 3 page suivante). Tout comme le choix des méthodes de saisie, le choix de la méthode de codage doit résulter d'une analyse coût / bénéfice / qualité / délais / faisabilité. Chacune des méthodes disposent d'avantages et d'inconvénients qui doivent être analysés.

II.2 II.2 RECOMMANDATIONS LIEES AUX METHODES DE CODAGE

- analyser les inconvénients et les avantages de chaque méthode de codage, les risques et les gains en fonction du contexte national pour déterminer le choix de la méthode la plus adéquate ;
- aboutir à un choix participant au renforcement des capacités des systèmes statistiques nationaux, et plus précisément des cadres nationaux et des institutions en charge de l'opération ;

Dans le cadre du contexte des pays de la sous région et de l'amélioration de la qualité des données, AFRISTAT recommande l'utilisation du codage assisté par ordinateur. Cependant, certains pays disposent des acquis et ressources pouvant justifier l'utilisation du codage automatique.

III. Méthode d'apurement[3] des données

L'évolution des technologies permet aux institutions en charge du recensement de la population et de l'habitat d'utiliser des méthodes de plus en plus complexes pour l'apurement des données (corrections et imputations). La stratégie d'apurement des données fait partie de l'ensemble du processus de la phase d'exploitation et de traitement des données. Pour cette raison, il importe de choisir cette dernière dans la phase de préparation de l'opération.

[2] Les dictionnaires de codifications son remis à jour au cours du dépouillement.
[3] Nous entendons ici par apurement, les méthodes de corrections et d'imputation

TABLEAU 3 : METHODES DE CODAGE

METHODES	***AVANTAGES***	***INCONVENIENTS***	***OBSERVATIONS/REMARQUES***
Codage manuel	Le codage s'effectue préalablement à la saisie	Travail fastidieux	Il s'agit de la méthode la plus employée dans les pays de la sous région. L'agent compare la réponse collectée et la traduit par un code grâce à un dictionnaire de codification
		Donne lieu à des erreurs plus importantes et plus nombreuses que les autres méthodes de saisie	
Codage assisté par ordinateur	Le système étant informatique, il peut intégrer un plus grand nombre de règles pour aboutir au codage et permet d'améliorer la qualité des données (réduction du nombre d'erreur)	Nécessite un temps de préparation important	Cette méthode s'appuie sur l'utilisation des systèmes informatiques. En réalité, les dictionnaires de codification utilisés par l'agent de codage sont informatisés. L'opérateur de saisie devient le codeur et choisit lors de la saisie l'entrée dans le dictionnaire de codification et le code est automatiquement généré.
	Permet un gain de temps (réduction du temps de saisie). Le dictionnaire de codage peut prendre en compte les potentiels d'erreur de saisie en offrant la possibilité de confirmation de l'intitulé exact (bloul… boulanger).	Nécessite de prendre en compte le coût	
	Méthode éprouvée par de nombreuses expériences qui montrent l'efficacité de la méthode)		
Codage automatique		Nécessité de prendre des précautions car les erreurs soit dans les algorithmes, soit dans les dictionnaires de codification peuvent aboutir à de mauvais codage	Le codage automatique se fait sans intervention humaine et fait appel à des algorithmes complexes pour comparer les dictionnaires de codifications et les réponses contenues dans les questionnaires
		Système complexe qui nécessite un investissement d'appropriation	
		Proportion de concordances faible pour des codages complexes comme l'activité professionnelle…	
		Impossibilité de coder automatiquement l'ensemble des réponses et nécessité de retour au codage assisté par ordinateur ou au codage manuel pour certaines variables	

III.1 AVANTAGES ET INCONVENIENTS DE CHAQUE METHODE

Dans le domaine des corrections, c'est-à-dire du contrôle de validité et de cohérence (méthode de vérification descendante, de vérification portant sur plusieurs variables…) et des imputations (cold deck, hot deck, ventilation et/ou répartition au prorata, méthode automatisée d'imputation dynamique) les techniques évoluent également rapidement. Ainsi, la méthode qui consiste à imputer des données, c'est-à-dire combler des lacunes à partir des autres informations peut se réaliser à travers différentes méthodes.

Chaque méthode devrait être rigoureusement testée avant sa mise en œuvre effective. Les aspects de corrections et d'imputations des données concernent différents aspects et en particulier l'étendue des corrections (en fonction du contexte). D'autre part, les manuels des Nations Unies[4] rappellent que les expériences des pays montrent que les systèmes d'apurement des données les plus complexes et donc les plus difficiles et les plus coûteux, peuvent être la source de nouvelles incohérences de données. Ainsi, il est parfois plus utile d'adopter des procédures de correction simples en corrigeant les erreurs évidentes. En général, on estime que les avantages d'une méthodologie de corrections et imputations minimaliste et réaliste effectué sur les variables démographiques de base dépassent largement les avantages éventuels des méthodes plus complexes comme la méthode hot deck.

III.2 RECOMMANDATIONS LIEES AUX METHODES DE CORRECTION

- adopter un système de correction relativement simple qui consiste à corriger les erreurs évidentes et qui réduit les risques et les coûts associés aux systèmes complexes d'apurement ;
- favoriser l'utilisation d'un système de correction automatique car les procédures manuelles sont inefficaces et n'améliore pas les résultats. D'autre part il est difficile de retracer les corrections effectuées ;
- promouvoir le travail en synergie entre informaticiens (programmateurs) et statisticiens (analystes) ;
- harmoniser les systèmes de corrections avec les expériences de l'institut de statistique pour assurer une comparabilité entre les sources (enquêtes et derniers RGPH) ;
- adopter un système de correction des données qui soit en harmonie avec les normes des Nations Unies (DSNU) ou le l'Organisation internationale du travail (OIT) ;

---

[4] **Division de la Statistique des Nations Unies**, « *Manuel d'organisation des recensements de la population et de l'habitation, Etudes méthodologiques* », Département des affaires économiques et sociales, Séries F n°83, ST/ESA/STAT/SER.F83, New York, 2002, 251p.

Division de la Statistique des Nations Unies*, Rapport préliminaire de l'atelier des Nations Unies sur les principes et recommandations pour les Recensements Généraux de la Population et de l'Habitat (RGPH) de la décennie 2010*, Bamako 6 au 9 novembre 2006.

Division de la Statistique des Nations Unies, « *Manuel de vérification des recensements de la population et de l'habitat - Etudes méthodologiques* », Département des affaires économiques et sociales, Séries F n°82, ST/ESA/STAT/SER.F82, New York, 2002, 125p

- effectuer un compte rendu ou rapport de l'ensemble des procédures de corrections effectuées et archiver ce rapport pour les prochaines opérations (métadonnées).

IV. Autres aspects du traitement des données

Indépendamment des méthodes de saisie, de codage et d'apurement des données, il est important de prévoir la gestion des données et plus spécifiquement le stockage et la sécurisation des données.

IV.1 STOCKAGE ET ARCHIVAGE

Les expériences de certains pays de la sous région montre que l'archivage des données est parfois problématique. Certains pays ont parfois perdu les bases de données de départ (bases des données brutes), voir des parties de fichiers de données ou la totalité. D'autre part, de part le fait que le processus de traitement de l'information s'effectue sur différents ordinateurs, à différents endroits géographiques dans le cas de mise en place de centre régionaux de saisie, il est nécessaire de prévoir bien en amont le schéma de gestion des données. La procédure de stockage dépend de l'organisation mise en place et des méthodes de traitement, de saisie, d'apurement des données (volume des données et complexité du système d'exploitation). D'autre part, les expériences montre la persistance de virus informatiques qui nécessite de prévoir ces aspects.

IV.2 RECOMMANDATIONS LIEES AUX METHODES DE CORRECTION

- prévoir l'élaboration d'un plan de gestion des données en amont et un plan de communication (entre les équipes) de l'opération de traitement des données ;
- assurer la conservation des copies et de chacune des versions des fichiers de données tout au long du processus de traitement ;
- assurer la constitution de doubles des données (au sein de l'institution et hors de l'institut) en cas de catastrophes majeures ;
- prévoir une stratégie de récupération en cas de perte de données, d'altération d'un fichier, de problème de serveur ;
- adopter un système et un réseau de protection des données avec un accès limité au personnel concerné (anti-virus y compris) ;

CONCLUSION

Le RGPH constitue une opération complexe qui nécessite une mobilisation accrue de temps et de moyens (humains, financiers et matériels). Le choix des outils pour la mise en œuvre des différentes activités de l'opération, telle que les méthodes de saisie, de codage et d'apurement des données doit résulter d'une analyse raisonnée. En effet, ce choix aura directement un impact sur les résultats et sur l'atteinte des objectifs de l'opération et le respect des délais. A cet effet, ces choix devraient être effectués bien en amont du début de l'opération en terme d'analyse coût / bénéfice / qualité / délais / faisabilité.

Annexe 1 : Calendrier des recensement des pays membres d'AFRISTAT

Périodes et dates des RGPH	1985 - 1994	1995 - 2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Benin	1992	2002								2012 (S)		
Burkina Faso	1985 (B)	1996		2006								
Burundi	1990	--				2008(S)						
Cameroun	1987	--	2005 (D)									
Comores	1991	2003									2013 (S)	
Congo	1994 (G)	1996		2006(S)								
Cote d'Ivoire	1988	1998				2008(P)						
Gabon	1993 (B)	2003									2013 (S)	
Guinée	--	1996				2008(S)						
Guinée Equatoriale	1994 (B)	2002										
Guinée-Bissao	1991	2001				2008(P)						
Mali	1987	1998	--				2009 (S)					
Mauritanie	1988	2001	--						2011 (S)			
Niger	1988	2001							2011 (S)			
République Centrafricaine	1988	2003									2013 (S)	
Senegal	1988	2002								2012 (S)		
Tchad	1993	--				2008 (S)						
Togo	--	--					2009 (S)					

(A) Recensement administratif
(B) Recensement de la population uniquement
(D) Pas de diffusion d'informations claires sur la programmation
(S) Programmé mais pas encore réalisé
(P) Estimation de la date du recensement effectuée selon le rythme décennal

Annexe 2 : Mode de comptage : Population présente, population résidente habituellement, intention de résider ou non…

Le choix de la méthode de comptage doit être mis en parallèle avec l'objectif du RGPH : fournir une base fiable par le comptage précis d'une population à un moment donné indispensable à la planification efficace, à la prévision des besoins en services, à la distribution des ressources, à la définition des découpages pour la mise en œuvre de la représentation électorale et politique.

Les pays sont principalement intéressés par le comptage et la répartition des résidents habituels car le lieu de résidence habituelle est le meilleur indicateur de l'endroit où les personnes ont besoins et consomment des services. Ainsi, le nombre des résidents habituels est l'indicateur le plus fiable en termes de planification et de politique. L'utilisation du mode de comptage « des résidants habituels » identique dans la sous région et dans les pays membres d'AFRISTAT nécessite de prendre des précautions. Différents mode de comptage existent : population présente, population résidente habituelle ou population liée aux services. Chaque méthode de comptage dispose d'avantages et d'inconvénients.

La population présente *correspond aux individus recensés à l'endroit où ils se trouvent lors du dénombrement (logement où ils passent la nuit précédent le recensement). Ainsi, la population présente comporte les résidents étrangers qui se trouvent dans le pays lors du recensement et exclue les résidents habituels du pays qui sont absents au moment du dénombrement. Le choix de recenser la population présente évite certaines complications liées à l'application de la notion de lieu de résidence habituelle et réduit les biais d'erreur liés au double comptage. Mis à part les avantages liés à sa simplicité, le comptage de la population présente est moins coûteux : le recensement ne doit pas collecter d'informations supplémentaires sur les résidents habituels qui ne sont pas à leur résidence habituelle. Cependant, il ne fournit pas un comptage exhaustif des résidents habituels et par conséquent ne donne pas la véritable répartition géographique des résidents habituels recherchée à des fins de planification et de politiques efficaces.*

La sous région se caractérise par de fortes migrations (que ces dernières soient internes ou internationales) et par conséquent une part importante de la population ne se trouve pas sur le lieu de résidence habituelle lors du recensement. Ensuite les caractéristiques des résidents habituels absents peuvent être souvent complètement différentes de celles des non-résidents présents (migration de travail, migration de confiage par exemple). Enfin, lorsque l'objectif principal d'un recensement est d'offrir une base de sondage de qualité, il est important de disposer des informations précises sur les ménages et leur composition (pour les futurs pondérations des enquêtes) et le comptage de la population présente ne permet pas de recenser les personnes absentes et habituellement présentes dans leur ménage. Sur la majorité des pays de la zone, on utilise **le comptage de la population résidente habituelle***. Cette façon de dénombrer les individus donne des informations de meilleure qualité pour les questions de planification et de politique axées sur la demande de services qui concerne les ménages, les familles et les migrations intérieures.*

En 1981, **le RGPH du Togo** *défini comme lieu de résidence la localité où une personne habite la plupart du temps, même si elle a l'habitude de partir périodiquement dans une autre localité : « Est résident dans une localité donnée tout individu qui y habite depuis (6) mois au moins »*[5]*. Lors de ce recensement, on distingue les résidents présents (tout résident qui se trouve à son lieu de résidence habituelle lors du passage de l'agent recenseur) et/ou ayant passé la nuit précédente à son lieu de résidence habituelle. Un résident présent est aussi un individu ayant passé la nuit précédente au logement habituel sans pour autant être présent au moment du passage de l'agent recenseur. Enfin, on dénombre les résidents absents, c'est-à-dire tout résident n'ayant pas passé la nuit précédant le passage de l'enquêteur à son lieu de résidence habituelle et qui ne s'y trouve pas au moment du RGPH. Dans le cadre du RGPH de 1981 au Togo, tout individu absent depuis 6 mois ou plus n'est pas recensé.* **Le RGPH du Bénin** *(2002) distingue également la population résidente (résidents présents et résident absents) et la population de fait (visiteurs). Cependant, les documents en notre possession ne donnent pas la définition exacte de résident (durée de résidence habituelle)*[6]*.* **Le**

---

[5] Manuel d'instructions aux enquêteur 1981, Togo.

[6] Synthèse des résultats du troisième recensement général de la population et de l'habitat, Février 2002, Direction des Etudes Démographiques, Cotonou, décembre 2003, 38p.
Analyses des Résultats du RGPH 3, Tome 1, « Répartition spatiale, Structure pas sexe et âge et Migration de la Population au Bénin », Direction des Etudes Démographiques, Cotonou, Octobre 2003, Institut national de la Statistique et de l'Analyse Economique

Burkina Faso (RGPH 2006) *utilise également la méthode de comptage des résidents habituels.* **Le RGPH de 1998 de la Côte d'Ivoire** *bien qu'utilisant la même méthode de comptage ajoute l'intention de rester plus de 6 mois : « Est considéré comme résident toute personne qui vit dans son lieu de résidence de façon habituelle ; c'est-à-dire depuis au moins 6 mois ou encore depuis moins de 6 mois mais avec l'intention d'y rester plus de 6 mois »*[7]*.* **Le RGPH 2002 du Sénégal** *définit également la résidence en fonction d'une nuit de référence au moment du recensement et d'une durée de présence de 6 mois au moins. En revanche le visiteur est défini comme la personne ne vivant pas habituellement dans le ménage mais s'y trouvant depuis moins de 6 mois*[8]*. Ici le visiteur selon le RGPH du Sénégal peut être résident pour le RGPH de Côte d'Ivoire (si le visiteur a l'intention de résider plus de 6 mois).*

Selon les définitions retenues par les Nations Unies, les résidents habituels peuvent avoir la nationalité, ou ne pas l'avoir, et peuvent également comprendre des personnes sans papiers, etc., des demandeurs d'asile ou des réfugiés. Les résidents habituels peuvent aussi comprendre des étrangers qui résident ou ont l'intention de résider de manière permanente sur le territoire, soit pendant la majeure partie des 12 derniers mois, soit depuis 12 mois, voire davantage, selon la définition du lieu de résidence habituelle adoptée par le pays. Sur les pays de la sous région, la notion de « résidant habituel » prend souvent comme période de référence les 6 derniers mois et parfois moins de 6 mois avec l'intention d'y résider plus de 6 mois.

Pour produire un comptage de la population résidente habituelle, des renseignements sont requis sur tous les résidents habituels, ainsi que sur l'adresse de leur lieu de résidence habituelle qu'ils soient présents ou non lors du recensement, de manière à garantir un comptage exhaustif. Il peut y avoir des biais d'erreurs dans les informations concernant les résidents habituels qui sont absents du pays lors du comptage surtout si aucune personne ne peut fournir des renseignements à leur sujet. D'autre part, certaines personnes considèrent ne pas avoir de lieu de résidence habituelle (nomades).

Il existe un troisième mode de comptage de la population (liées aux services). Ce dernier devient nécessaire si les deux autres modes de comptages ne reflètent pas fidèlement la demande ou la fourniture de services dans une région spécifique (zone de population de résidents étrangers qui franchissent régulièrement la frontière pour travailler ou consommer des services). Dans le cadre du comptage de la population liée aux services, outre l'estimation du nombre de résidents habituels, il est nécessaire de collecter des informations concernant les lieux où les individus « travaillent » ou « consomment » des services.

Comte tenu des expériences des pays de la région et des définitions adoptées dans les recensements précédents, les prochains RGPH devront considérer la population habituellement résidente. La notion d'intention de résider plus de 6 mois lorsqu'un individu est présent depuis moins de 6 mois mérite d'être discutée. D'un côté, cela entraîne une complication de la définition et de la compréhension du mode de comptage des « résidents habituels ». D'un autre côté, écarter cette notion entraîne la non possibilité de comparaison de l'évolution de la population effective de la Côte d'Ivoire entre 1998 et le RGPH suivant.

---

[7] Recensement général de la population et de l'habitat de 1998, RGPH – 98, Volume IV : Analyses des résultats, Tome 1 : Etat et Structure de la population, Essoh Badou, Démographe, Abidjan, juin 2001, 118p, page 13.

[8] Résultats du troisième recensement général de la population et de l'habitat (2002), Rapport National de présentation, décembre 2006, 130p.